

# Archival Legacy Investigations of Circumstellar Environments: Overview and First Results

Élodie Choquet<sup>a</sup>, Laurent Pueyo<sup>a</sup>, J. Brendan Hagan<sup>a,b</sup>, Elena Gofas-Salas<sup>a,f</sup>, Abhijith Rajan<sup>g</sup>, Christine Chen<sup>a</sup>, Marshall D. Perrin<sup>a</sup>, John Debes<sup>a</sup>, David Golimowski<sup>a</sup>, Dean C. Hines<sup>a</sup>, Mamadou N'Diaye<sup>a</sup>, Glenn Schneider<sup>c</sup>, Dimitri Mawet<sup>e</sup>, Christian Marois<sup>d</sup>, Rémi Soummer<sup>a,†</sup>

<sup>a</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218, USA;

<sup>b</sup>Purdue University, West Lafayette, IN 47907, USA;

<sup>c</sup>Steward Observatory, University of Arizona, North Cherry Avenue, Tucson AZ 85721, USA;

<sup>d</sup>NRC Herzberg Institute of Astrophysics, West Saanich Road, Victoria, BC V9E 2E7, Canada;

<sup>e</sup>ESO, Alonso de Cordova 3107, Casilla 19001, Santiago 19, Chile;

<sup>f</sup>Institut d'Optique Graduate School, Palaiseau Saint-Étienne and Bordeaux, France;

<sup>g</sup>Arizona State University, Phoenix, AZ 85004, USA;

## ABSTRACT

We are currently conducting a comprehensive and consistent re-processing of archival HST-NICMOS coronagraphic surveys using advanced PSF subtraction methods, entitled the Archival Legacy Investigations of Circumstellar Environments program (ALICE, HST/AR 12652). This virtual campaign of about 400 targets has already produced numerous new detections of previously unidentified point sources and circumstellar structures. We present five newly spatially resolved debris disks revealed in scattered light by our analysis of the archival data. These images provide new views of material around young solar-type stars at ages corresponding to the period of terrestrial planet formation in our solar system. We have also detected several new candidate substellar companions, for which there are ongoing followup campaigns (HST/WFC3 and VLT/SINFONI in ADI mode). Since the methods developed as part of ALICE are directly applicable to future missions (JWST, AFTA coronagraph) we emphasize the importance of devising optimal PSF subtraction methods for upcoming coronagraphic imaging missions. We describe efforts in defining direct imaging high-level science products (HLSP) standards that can be applicable to other coronagraphic campaigns, including ground-based (e.g., Gemini Planet Imager), and future space instruments (e.g., JWST). ALICE will deliver a first release of HLSPs to the community through the MAST archive at STScI in 2014.

**Keywords:** HST, NICMOS, Coronagraphy, Post-processing, debris disks, exoplanets

## 1. INTRODUCTION

Since the detection of the first exoplanet,<sup>1</sup> our understanding of planetary systems and their formation has been revolutionized by the discovery of more than a thousand other extra-solar planets. Radial velocity and transit measurements have been the most successful methods, with hundreds of discoveries through several surveys over the past 15 years.<sup>2–4</sup> However, these methods are limited to the detection of planets relatively close to their host stars, and need very long time-scales to confirm planets on very wide orbits. Complementary, direct imaging provides observations of the whole environment of nearby stars at a glance, enabling detection and characterization of dusty disk and of systems with massive planets both spatially and spectrally.<sup>5–8</sup>

However, such observations are very challenging due to the large contrast between the circumstellar objects and their host star, when closer than 1". Despite multiple surveys conducted with telescopes equipped with coronagraphic systems, only a handful of bright debris disk have been imaged until the late 2000's,<sup>9</sup> and the first images of exoplanets were obtained only six years ago.<sup>10–12</sup> These discoveries were only possible thanks to careful star point spread function (PSF) subtraction during post-processing. At this time, the most common

---

Correspondence should be sent to choquet@stsci.edu, †: Principal Investigator

PSF subtraction method merely consisted in subtracting the image of a reference star, or a median image of a collection of reference stars from the target images. However, this method is still not optimal to reveal faint structures less than 2'' from the host star, due to slight variations of the wavefront (e.g. thermal breathing for space telescopes, residual errors from atmospheric turbulence for ground based telescopes with adaptive optics). For instance, in the case of the Hubble Space Telescope (HST), well-known for its stability,<sup>13</sup> small temporal variations<sup>14</sup> (focus breathing and cold mask shifts induced by periodic thermal variations along an HST orbit) affect classical PSF subtraction methods down to a level that prevents coronagraphic imaging to reach the fundamental limit of the speckle's photon noise within the inner  $\sim 2''$  of the PSF.

New post-processing methods based on the use of large libraries of reference stars have opened a new era in high-contrast exploration of circumstellar environments. The reference library is here used as a sample of optical path random realizations to generate a synthetic PSF as close to science image star PSF as possible. Since the development of the *Locally Optimized Combination of Images* (LOCI) algorithm,<sup>15</sup> and more recently with the use of Principal Component Analysis algorithms,<sup>16,17</sup> about 40 substellar companions have been imaged around nearby stars\*.

These post-processing methods have now become standard techniques for high contrast imaging. However, numerous coronagraphic programs have been carried out before 2010 without such optimized processing, and thus these programs potentially have unseen circumstellar structures that are within reach of the contrast achieved with these recent techniques. It has actually been demonstrated with the re-discovery of the HR 8799 planetary system<sup>5,18</sup> using HST NICMOS archival data acquired 10 years before their first images.<sup>10</sup>

In this context, we initiated a consistent and comprehensive re-processing of the entire NICMOS coronagraphic archive using advanced post-processing techniques, in order to reveal hidden objects and reach better detection limits with state-of-the-art algorithms. With about 400 targets in the archive, large reference libraries can be easily assembled to optimize the computation of the most accurate synthetic PSF for each science image. In this paper, we present an overview of this project, entitled *Archival Legacy Investigation of Circumstellar Environment*<sup>†</sup> (ALICE). We describe the NICMOS archive content in Sec. 2, then we detail the architecture of the ALICE pipeline in Sec. 3. First results of the project are presented in Sec. 4. We then conclude by presenting the prospects of the ALICE project in Sec. 5.

## 2. NICMOS ARCHIVE CONTENT

Our database contains all the NICMOS images which have been reprocessed through the HST archival LAPLACE program<sup>‡</sup> (*Legacy Archive PSF Library And Circumstellar Environments*),<sup>19</sup> i.e. all NICMOS coronagraphic images (acquired with the NIC2-CORON aperture) from cycle 7 through cycle 15, and the additional program 11155 (cycle 16, PI: M. Perrin). The LAPLACE reprocessing consists of an improved recalibration of the NICMOS coronagraphic data using contemporary flat field frames, observed dark frames and a better bad-pixel correction. This recalibration is crucial to build optimal reference libraries needed for advanced PSF subtraction algorithms as will be described in Sec. 3. The additional cycle 16 program has been added to our database to analyze the interface of the ALICE pipeline with non-LAPLACE data. Although included in the following description of our database, none of the results presented in this paper are obtained with data from this program.

The database contains images from 405 stars. Their spectral type distribution is roughly equally balanced between types M, K, G, and other spectral types, for stars with identified spectral type<sup>§</sup> (see Fig. 1, left). About 50 % of the stars with known parallax<sup>¶</sup> in our target list are within 35 pc from Earth (see Fig. 1, right).

The images in our database are distributed into 40 programs (See Fig. 2). 16 of them have been carried out between 1997 and 1999, during the first era of NICMOS during which the instrument was cooled with a nitrogen-ice cryogenic dewar. The cooling system was replaced by a cryogenic neon cryocooler in 2002 during Hubble Service Mission 3B, after the dewar ran out of nitrogen in 1999. The last 24 programs of our database

---

\*Source: The Extrasolar Planets Encyclopedia, <http://exoplanet.eu>

<sup>†</sup>HST-AR-12652, PI: R. Soummer

<sup>‡</sup>HST-AR-11279, PI: G. Schneider

<sup>§</sup>Source: The Simbad astronomical database, <http://simbad.u-strasbg.fr/simbad/>

<sup>¶</sup>Source: The Simbad astronomical database, <http://simbad.u-strasbg.fr/simbad/>

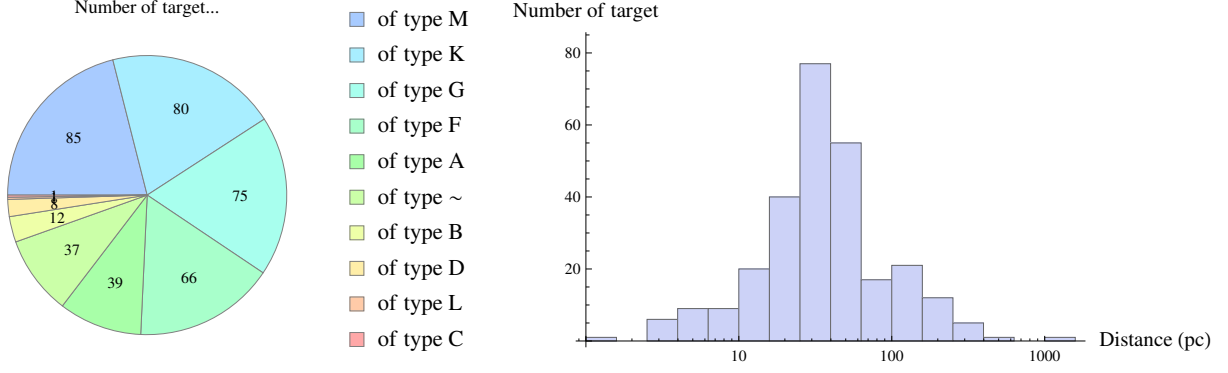


Figure 1. **Left:** Distribution of spectral types of our target list. **Right:** Histogram of the known distances of the stars in our database. 131 targets have unreferenced parallax measurements, out of the 405 stars in our target-list. Source: the SIMBAD database, <http://simbad.u-strasbg.fr/simbad>.

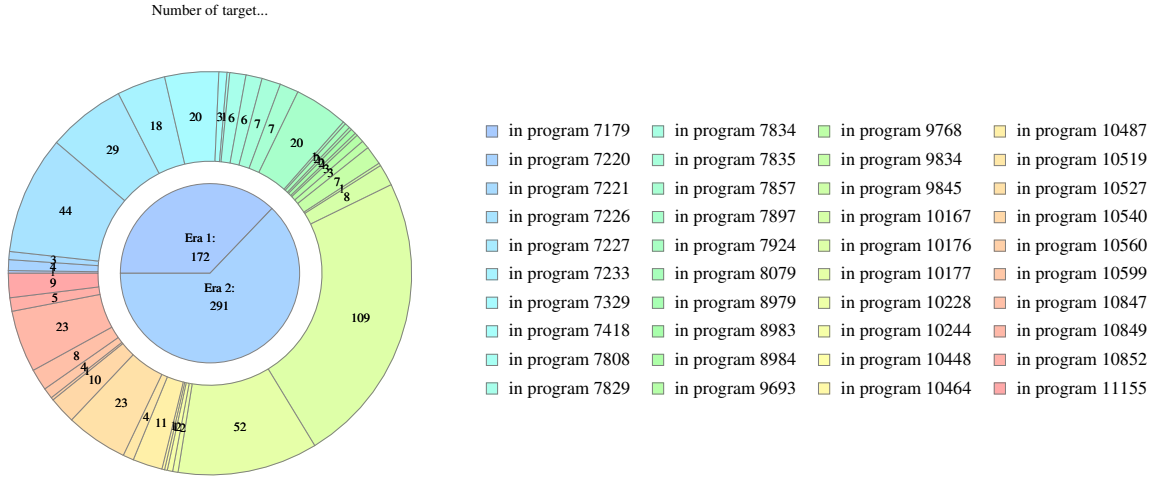


Figure 2. Number of target per NICMOS coronagraphic program. The programs from 7179 to 8079 were conducted during the first Era of NICMOS. The programs from 8979 to 11155 were conducted between 2002 and 2008, after replacement of the nitrogen-cooler by a cryogenic neon cryocooler. All these data have been reprocessed as part of the LAPLACE program, except program 11155 which have been recently added to our database to investigate on the compatibility between the ALICE pipeline and non-LAPLACE NICMOS data.

were conducted during this second era of the instrument, which represents 75 % of our database in number of images. This distinction between both eras is an important parameter in the image selection to assemble the reference libraries, as will be explained in Sec. 3.

Most of the images in our database are part of a survey program conducted to discover planets, debris disks, or protoplanetary disks. A brief description of these programs is provided in Table 1. Among the 405 stars in our database, 48 have been observed in several programs (12 % of our target list). If a companion candidate is detected around one of them, whether or not is it gravitationally bound to the parent star can potentially be confirmed using NICMOS data from another program, provided that the temporal baseline between both datasets and the star proper motion are important enough to differ the companion motion from a background object.

The vast majority of the images in our sample have been acquired with the J and H wide-band filter (67 % of the target list in the F160W filter, and 40 % in the F110W filter). The other filters were marginally used with

Table 1. NICMOS coronagraphic surveys conducted to search for planets and disks. Only surveys with more than 150 images are described.

Program	PI	# Target	NICMOS Filters	Type of survey
7226	E. Becklin	44	F110W, F160W, F165M, F180M	Planets (young and nearby)
7227	G. Schneider	29	F160W	Planets (around M stars)
7233	B. Smith	18	F110W, F160W, F204M, F237M	Disks (MS stars, IRAS excess)
7834	R. Rebolo	6	F110W, F160W, F180M, F207M	Planets (young, nearby, late)
10167	A. Weinberger	8	F171M, F180M, F204M, F222M	Disks (known disks)
10176	I. Song	109	F160W	Planets (young nearby)
10177	G. Schneider	52	F110W, F160W	Disks (IRAS excess)
	D. Ardila	11	F110W	Disks (Beta pic moving group)
10519	J. Simpson	4	POL0L, POL120L, POL240L	Disks (YSO, massive)
10527	D. Hines	23	F110W	Disks (Spitzer excess)
10540	A. Weinberger	10	F110W	Disks (IR excess)
10847	D. Hines	8	POL0L, POL120L, POL240L	Disks (known disks)
10849	S. Metchev	23	F110W	Disks (Spitzer excess)
11155	M. Perrin	9	F110W, POL0L, POL120L, POL240L	Disks (Herbig Ae)

the coronagraphic mode of NICMOS. Yet, 95 targets (23 % of the target list) have been observed at least two different filters. The comprehensive reprocessing of these targets will thus provide color measurements of any object detected around the host star (co-moving companion, disk, as well as background object).

Although the majority of our sample has been observed with several different orientations of the telescope (82 % of them in exactly two orients of HST), 82 stars have been only observed for a single orientation of the telescope, which significantly increase the false-alarm probability of faint point source detection.

### 3. THE ALICE PIPELINE

The ALICE pipeline has been designed to enable processing by several users of any coding level. It is entirely developed in *Wolfram Language* with *Mathematica* and utilizes several graphical user interfaces (GUIs) to ease its exploitation. Several tools have been developed to keep tracks of the user activities and of the intermediate files generated.

The pipeline is structured into three main sections:

- **Data preparation:** In this part, images from several programs are gathered, aligned, and sorted to create reference image cubes that are necessary for advanced PSF subtraction algorithms.
- **Data reduction:** The targets to be reduced are selected, a parameter space is defined to optimize the reduction, then the selected images are reduced using the KLIP algorithm.<sup>16</sup>
- **Data analysis:** All reduced images of a target are examined and combined, and the reduction parameter space is explored to determine the optimal reduced image. Residual speckles are analyzed to determine if they are real point sources or star residuals, then final science products are generated.

Several log files keep track of the operations performed on the data. The main log is a global workflow file which gives an overall view of the processing status of each program–target–filter combination. A log entry is added each time a target starts or ends being processed by one of the main section of the pipeline . Additional log entries are also generated at intermediate steps of the process. For each entry, the workflow indicates the user, pipeline version, and name of files generated or used at the different steps. The workflow is thus a key functionality for managing such a large program, including multi-user activities, or following the status of particular targets through each processing steps until final product delivery.

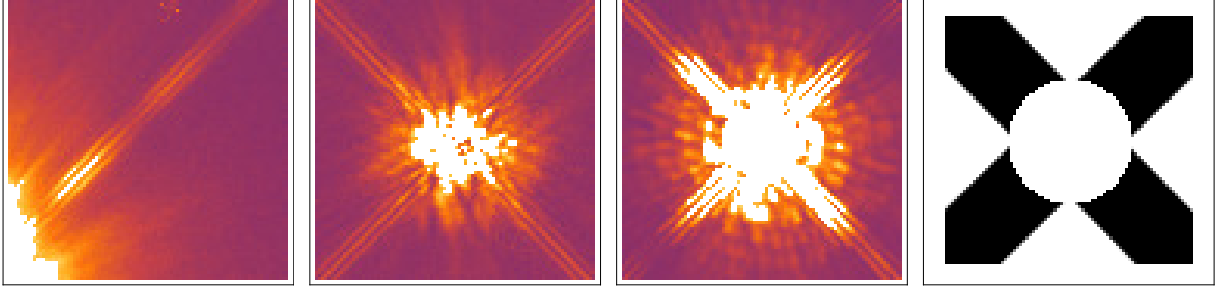


Figure 3. **Left**: typical bad acquisition image dismissed from the reference cube. **Mid-left**: potential good reference image. **mid-right**: F160W TinyTIM synthetic PSF, used as the initial reference for alignment. **Right**: Typical mask used for the alignment of the reference cube.

### 3.1 Construction of the Reference Libraries

The first part of the pipeline consists in building reference cubes that will be used by advanced PSF subtraction to reduce the science images. This is the most critical part of the pipeline, for a mis-prepared reference cube will result in a poor PSF subtraction at the next step of the pipeline. This part is also the most computation-time consuming.

#### 3.1.1 Image selection

The first step to select potential reference images by combing several archival programs, including the one containing the science target to be reduced. Ideally, one wants a large selection of PSFs similar to the target image to achieve an optimal subtraction, typically larger than 200 images. To achieve efficient PSF post-processing, two requirements have to be met:

- The selected programs have to have been conducted during identical NICMOS cooling eras, so either before or after 2002, but not a mix. Since the cooling system have been modified between the two eras, the detector responses (dark frames, bad pixels, flat fields) are not fully comparable despite similar reprocessing by the LAPLACE program.<sup>19</sup> References from one cooling era do not provide good matches to reduced a science image from the other instrument era.
- All the images in the reference cube have to be acquired with a unique and identical spectral filter. The instrument PSF is indeed wavelength-dependent, as well as the radial position of star-induced residual speckles. PSF subtraction from reference images with different filter systematically results in poor starlight suppression.

The selected images are then cropped to a squared field of view, hereafter the ALICE field of view (typically 6" across). The field of view is centered on the star position, which is identified using the location predicated from the LAPLACE processing (FITS file keywords "TARSIAFX" and "TARSIAFY").

Finally, the images are examined to flag and discard the obvious failed acquisitions. About 5 % of all images in our database has been flagged as a bad image. Fig. 3 presents examples of a bad acquisition image (left) and of a potentially good reference image (mid-left). Note that no further correction of bad pixel is applied to the images. Only a few bad pixels are occasionally present in the images thanks to the LAPLACE reprocessing, which do not significantly alter the performance of the data processing.

#### 3.1.2 Image alignment

The next step of a reference cube preparation is the fine alignment of all its images. This process is the most critical to achieve efficient PSF subtraction. The field centering on the star location predicted by the LAPLACE reprocessing is indeed not precise enough, because of instrumental variations during the spacecraft slew between the non-coronagraphic target acquisitions used to find the star centroid and the coronagraphic acquisitions.

The alignment is performed using a synthetic Tiny Tim PSF<sup>20</sup> simulated in the F160W filter, whose known position is at the center of the ALICE field of view. We define this as the absolute reference position for the alignment (see Fig. 3, mid-right). However, the model PSF is simulated without coronagraphic mask and is thus quite different from the cube images. We thus use a mask blocking out the core of the PSF and align the images only using the diffraction pattern of the telescope struts, which are clearly visible in NICMOS coronagraphic images (see Fig. 3, right, for a typical alignment mask). Since the unobscured Tiny Tim PSF is still slightly different from real images, we perform the alignment in two steps:

1. A median image from the reference cube is aligned with the Tiny Tim PSF;
2. All the images in the cube are then aligned with this aligned median image as reference.

The alignment mask is used in both steps for every image, to focus the alignment on the diffracted spiders and avoid mis-alignments in case of bright close binary or disk contamination. In addition, the images can be convolved by a gaussian filter (standard deviation chosen by the user), to account for the fact that NICMOS images are Nyquist-sampled at  $1.7 \mu\text{m}$ , and consequently undersampled at shorter wavelengths.

The default alignment algorithm used in the pipeline finds the position of the field of view maximizing the cross-correlation between an image and the reference image for alignment, filtered by the mask. All aligned images are then individually examined by the user to check the alignment with the reference image. Normalization of all the image of the cube is also performed to ease the examination of the alignment by the user.

Because of the third-order interpolation used for alignment, this process the most time-consuming of the pipeline. This method is particularly efficient for images in the F110W and F160W filter, where the diffraction pattern of the spiders is clearly visible. At longer wavelengths, the intensity of quasi-static speckles becomes dominant and impairs the alignment quality. Bad pixels can also occasionally impair the alignment when they spread over several pixels during the procedure because of sub-pixel interpolation. For these reasons, a careful examination of the alignment of all images is necessary at this step. We are currently analyzing different alignment algorithms to offer alternative solutions to the users of the ALICE pipeline depending on their need of celerity or robustness.<sup>21</sup>

### 3.1.3 Detection of bad references

The final step in the preparation of a reference cube is the identification of bad reference targets. To properly subtract a synthetic PSF with advanced algorithms, the PSF libraries must indeed be clean of any science components. Images in the reference cube with disks or point sources can indeed leave inconvenient artifacts in the final reduced images.

To detect bad references in a PSF library, we first reduce each of its images using the cube itself as the PSF library (removing all other images of the target being reduced to avoid self-subtraction). This helps reveals if companions or disks are hidden by the bright PSF. At this step, we use the LOCI algorithm<sup>15</sup> on a single zone, the full ALICE field of view (masking only the center of the image with a small mask of typical radius of  $\sim 5$  pixels), with a standard matrix regularization, identical for all the reduced images in the cube. For this first processing LOCI is preferable to PCA algorithms, because artifacts due to bad references always appear in negative in the reduced images, due to very aggressive subtraction of the LOCI algorithm. With PCA, the presence of a companion in a reference image appears in several Karhunen-Loève (KL) components of different ranks, with both negative and positive values (see Fig. 4). The projection of the science image on the KL basis also independently provides negative and positive coefficients, which can result in either negative or positive artifacts in the PSF subtraction, depending of the number of KL modes used to compute the synthetic PSF (See Fig. 5). Fig. 5 presents examples of PSF subtraction with the two algorithms, and shows that using PCA can be misleading to identify bad references, with positive artifacts that looks like faint companions.

The typical time to reduce one full image using LOCI with matrix regularization is  $\sim 180$  ms on a 2.8 GHz laptop. When all the images in the cube have been reduced, they are carefully examined to identify targets with circumstellar material, which are then flagged as bad references. The process (reduction, examination, identification of bad references) is repeated several time by using the reference cube cleaned of the bad references, to refine the reduction and detect fainter companions. Note that some of these bad references are also interesting science target, whose reduction will be refined and optimized in the next section of the pipeline.

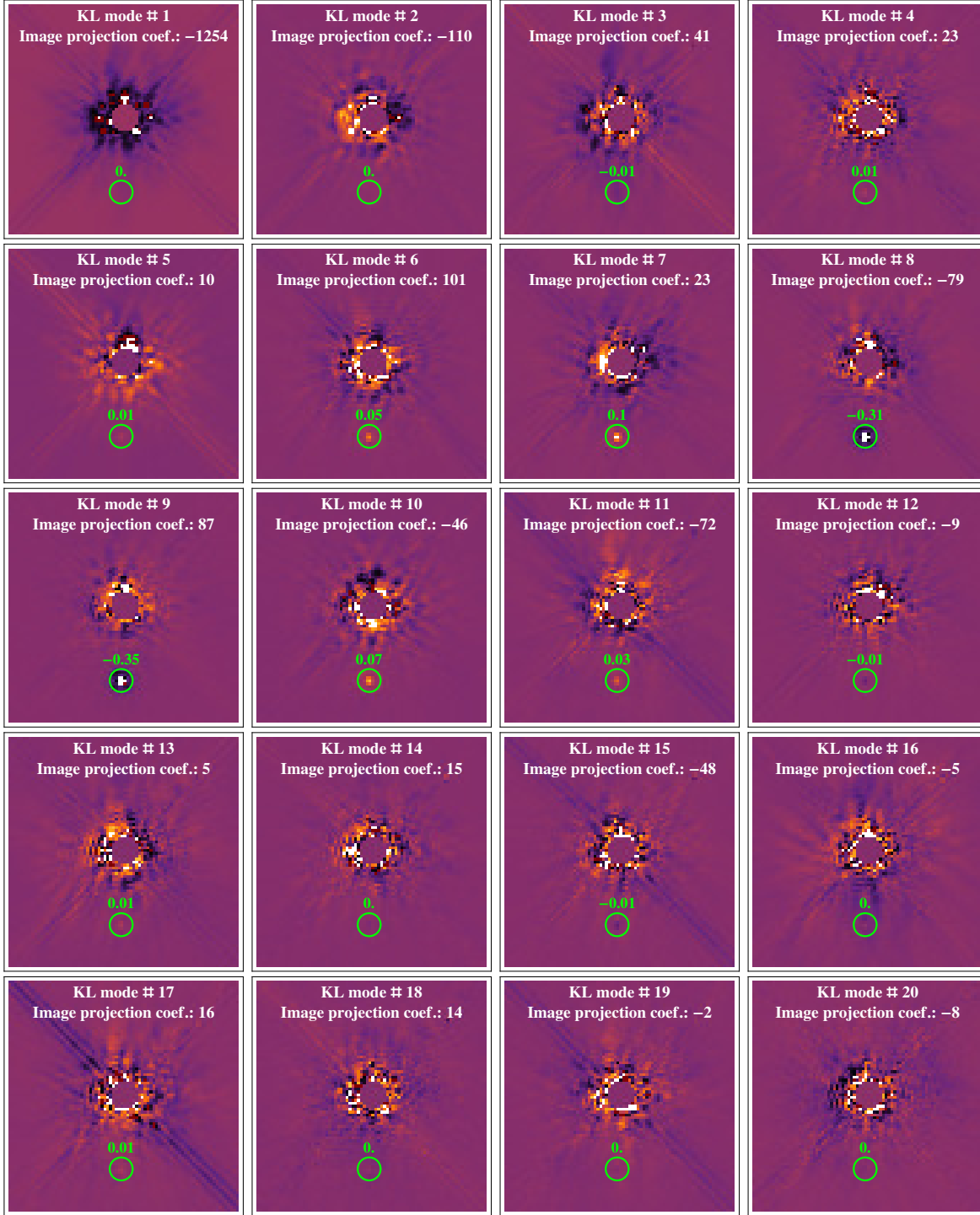


Figure 4. First twenty Karhunen-Loève modes of a toy reference cube. The cube contains the known bad reference TWA 5, which is a binary star.<sup>22</sup> The companion of the bad reference appears both with positive and negative values in the first KL modes of the cube, at positions indicated by green circles. The value above the circles are the pixel values of the modes at the center of the circles. The coefficients of the projection of a toy science image (known as being a good reference PSF) on this KL basis are displayed at the top of each image.

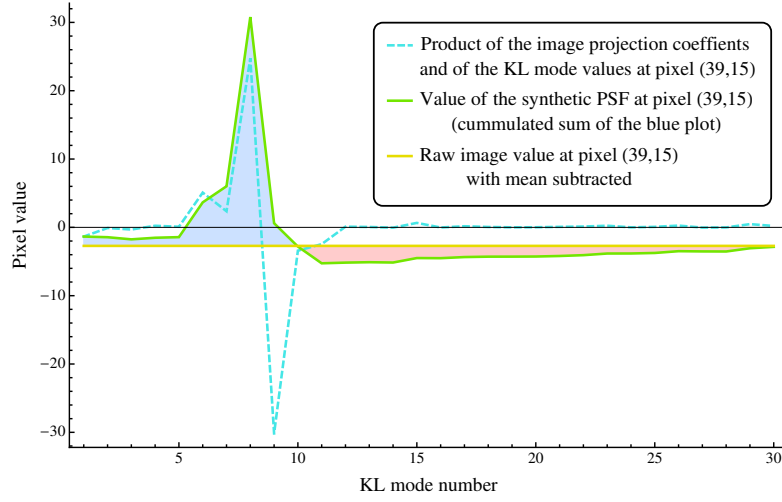


Figure 5. Details on the reduction of a toy science image with the KLIP algorithm,<sup>16</sup> using a reference cube with a bad target. The bad target is a binary star where the companion is a pixel position (39,15). The dashed blue curve shows the product between the KL mode value for this pixel (green values in Fig. 4) and the coefficient of the projection of the science image on the modes (white values in Fig. 4). The green curve is the cumulated sum of the later plot, and gives the value of the synthetic PSF provided by KLIP with increasing truncated mode value. The yellow curve shows the value of this pixel in the raw science image with mean subtracted. The difference between the two latter curves show that the use of this staled reference cube leads to the presence of a negative artifact (blue filling) for truncated values lower than 10 and to a positive artifact (red filling) for truncated values between 11 and 30, at the pixel position of the bad reference companion.

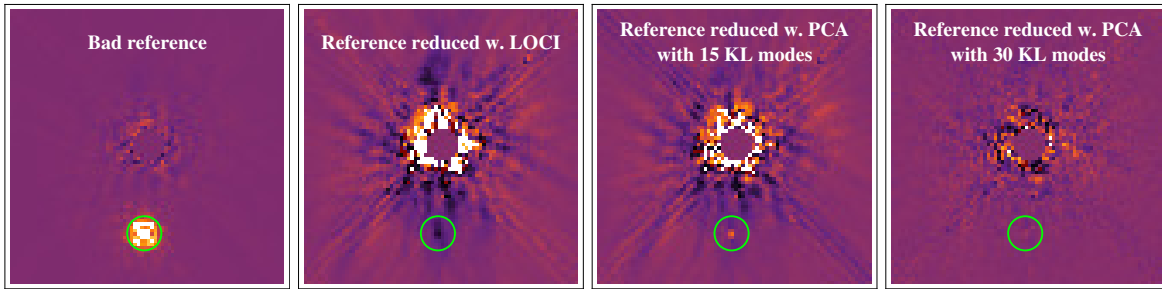


Figure 6. Impact of different reductions with a staled reference cube. **Left:** image of the bad reference in the PSF library. **Mid-Left:** science image reduced with LOCI, showing the presence of a negative artifact due to the bad reference in the PSF cube. **Mid-right:** science image reduced with the KLIP algorithm with 15 KL modes used to compute the synthetic PSF; the presence of the bad reference in the library adds a positive artifact which looks like a faint companion. **Right:** reduction with KLIP, with 30 KL modes to build the synthetic PSF; no artifact appears in this case.



### 3.2 Reduction

Once a reference cube has been assembled and carefully prepared, it can be used in the second main part of our pipeline, to reduce its science images. This part of the pipeline is program-oriented: once reduction parameters have been set by the user, all the images of a chosen program in the given reference cube are reduced altogether with the same parameters.

The ALICE pipeline has been designed to optimally process a large amount of data (the entire NICMOS coronagraphic archive is about 5600 images). This has motivated the choice of the KLIP algorithm<sup>16</sup> to reduce the data, which is based on principal component analysis and consists only of linear algebra operations. The typical computation time to reduce an image with KLIP is  $\sim 1$  ms.

This unique celerity enables the exploration of vast parameter spaces to find the optimal reduction for each target in reasonable amount of time. Our pipeline can explore seven different parameters:

**Mode:** The reference libraries can be optimized depending on the type of structure the user wants to detect. If a disk is being searched in a specific image, all the other images of this target have to be excluded from the reference library to avoid self-subtraction of the potential disk (especially for face-on disks). If planets are being searched, images of this target with sufficiently different telescope orientation to avoid self-subtraction can be used in the reference library. This is useful to include PSF images very similar to the target image being reduced. The ALICE pipeline supports both possible modes ("disk" and "planet") and optimizes the PSF library for each individual image being reduced.

**Zone:** The reduction process can be applied on specific zones of the science images instead of the full ALICE field of view. In the ALICE pipeline, portions of annuli centered on the star can be defined with four parameters (inner radius  $R_{min}$ , outer radius  $R_{max}$ , position angle of the center of the zone  $CenterPA$ , and zone azimuthal length  $\Delta PA$ ).

**Target preselection:** The reference libraries can be even more optimized to the individual images being reduced: with this parameter, we can change the size of the PSF library by selecting the reference images the most correlated with the science image. This selection might for instance automatically exclude stars with spectral types too different, and increase robustness to any other difference in the PSF (focus change, poor alignment etc.)

**Truncation number:** The last parameter defines the number of KL modes used to create the synthetic PSF subtracted to the science images. A parameter space with truncation numbers spanning the whole range of possible value is particularly useful to optimize the signal-to-noise ratio (SNR) on the reduced image and its possible circumstellar material, and maximize the contrast limits.

A GUI can be used to instantaneously visualize an image of the cube reduced with varying parameters to get a sense of the parameter space to be used for all the targets of the program.

Once a parameter space has been defined by the user, all the images of the selected program are reduced. The computation is parallelized by distributing the reduction of the different targets to the kernels of the machine and accelerate the process. For a medium size reference cube ( $\sim 450$  images), each science image is reduced with typically  $\sim 2250$  different parameter sets, in about 2.8 s. A three-dimension FITS cube is exported for each science image, with the reduced images for all the parameter sets, and can thus be analyzed later or by a different user.

### 3.3 Analysis

The final section of the ALICE pipeline is the analysis part. In this section, reduced image cubes of selected targets are re-imported and individually observed to find circumstellar material with a dedicated GUI (see Fig. 7).

For a given target being analyzed, all the images are rotated to have the North up, then are co-added and normalized to the total exposure time on the target. The reduction parameter space is manually explored by the user to find the optimal parameters for the combined image. The GUI proposes several tools to help the user choose the best parameter set:

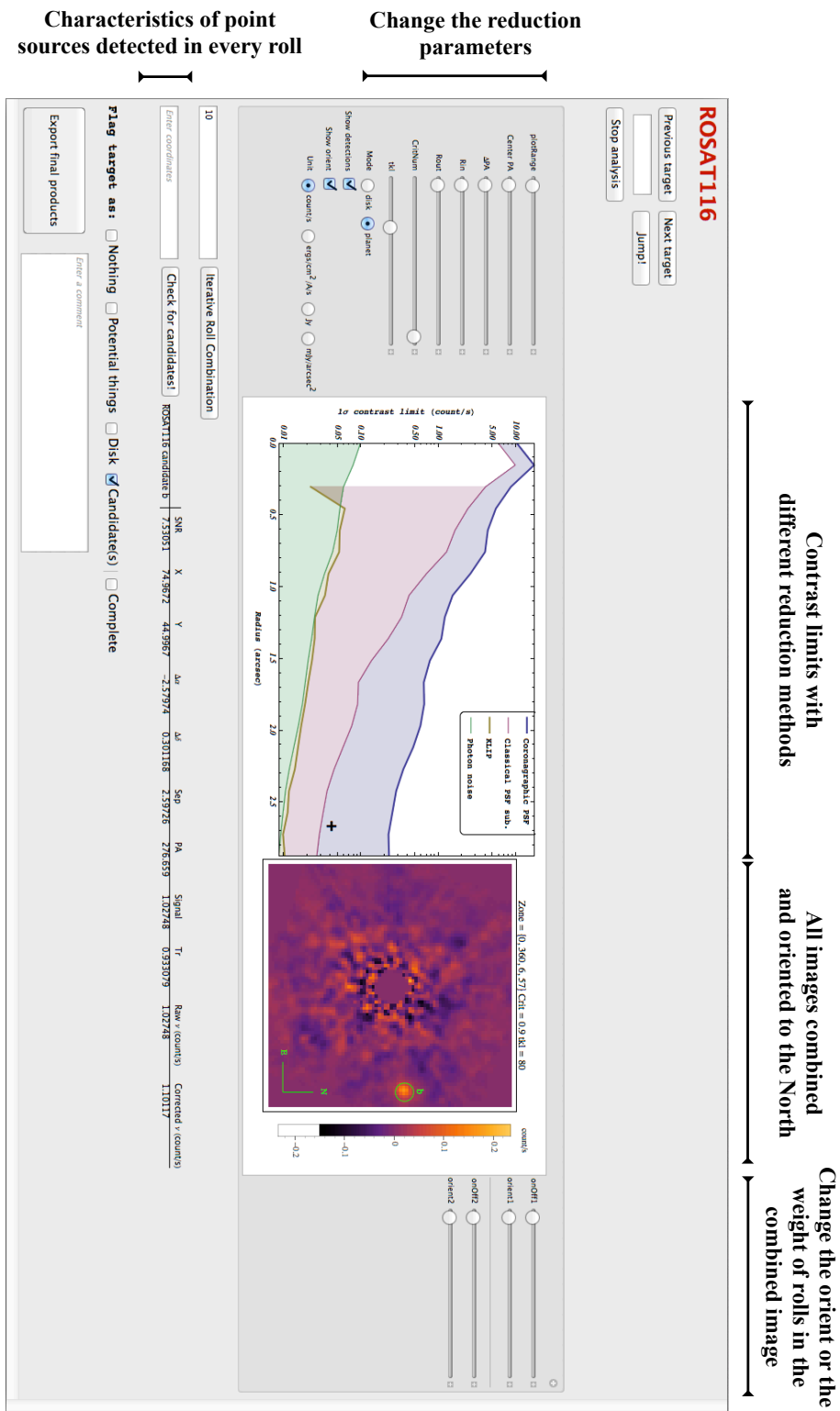


Figure 7. GUI used to for the analysis of a target (TWA 6 - ROSAT116 in this example).

- The user can change the orientation of a roll (combination of all images acquired with identical telescope orientation) or its weight in the displayed combined image. This enable to visually verify that a point source or a disk is in every roll, by checking how it rotates or decreases in intensity, respectively, in the combined image.
- Contrast curves are computed for four different combined images: the target raw combined image, the image processed classical PSF subtraction (median image of the reference cube), the combined image processed with KLIP, and the photon noise estimated from the raw images. The contrast curves are computed from the standard deviation in concentric annuli of width 2 pixels ( $\sim \lambda/D$ ).

Each individual reduced images can also be analyzed independently throughout the parameter space before combining them in the final image. If the some are not properly reduced (e.g. remaining bad pixel, misalignment), the user can exclude them from the combined image.

Once the optimal parameter set has been determined, the user can analyze and characterize residual point sources. If the target has been observed in several telescope orientations, detecting a companion at the same sky-coordinates in several rolls not only increases its SNR in the combined image but it also significantly decreases the false-alarm probability.

For each point source spotted by the user, the pipeline computes its precise astrometry and photometry in each roll, from the cross-correlation with a model PSF. Point sources with failed cross-correlation in at least one roll are considered as residual speckles and are discarded from the candidate list. Each candidate is then characterized in each individual reduced image, in each roll, and in the combined image, by computing:

- Its astrometry (both in pixel and sky coordinates relative to the star center).
- Its photometry, both in the reduced images and corrected from the algorithm throughput. The throughput is computed with forward modeling of the candidate, by computing the flux of a model PSF after subtraction by its projection onto the same KL vectors used for the reduced science image (see [16] for the analytic expression of forward modeling with the KLIP algorithm).
- Its SNR, by estimating the noise with the standard deviation in a small aperture of radius  $3 \lambda/D$  around the candidate and masking it out with an obstruction of radius  $\lambda/D$ .

Once the user is satisfied with the final image and the point source detections, high-level science products are exported, ready to be released to the MAST archive:

- Each individual reduced image in as many FITS files, with keywords detailing the pipeline parameters.
- The KL vectors and the zone mask for each image.
- A FITS cube with  $4 + N$  images with  $N$  the number of rolls. The  $N$  images are the individual rolls, combining all the reduced images acquired with the same telescope orientation. The four first images of the exported cube correspond to 1 and 2: the sum and difference of the two rolls both<sup>||</sup> rotated to the North respectively, 3 and 4: the sum and difference of these bins with their original orientations (See Fig. 8). If point sources or edge-on disks are present in the field of view, they appear at high SNR in the first image and self-subtract in the second, and appear twice with different orientations in the third and fourth images, respectively in positive-positive and positive-negative.
- A report with the reduced images and the candidates characteristics in the combined image, each individual roll, and each individual image.

---

<sup>||</sup>If the target has been observed with more than two telescope orientations, the roll images are rotated and combined to provide two binned rolls.

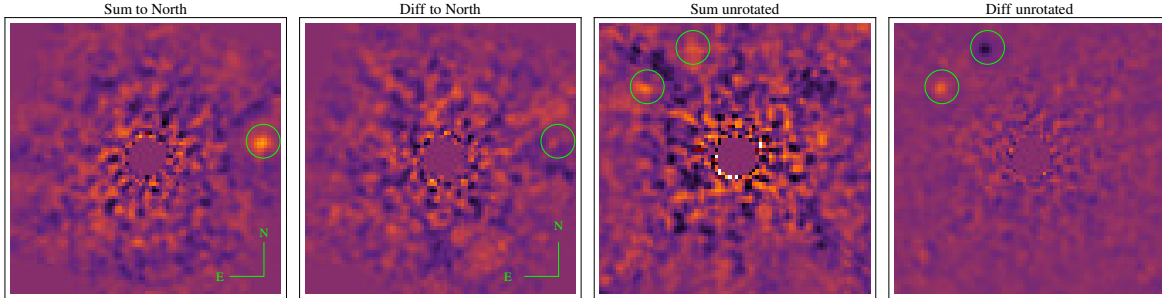


Figure 8. Combination of binned rolls exported in the FITS cube for the target TWA 6. **Left:** sum of the binned rolls oriented to the North; the candidate appears with a high SNR. **Center-left:** Difference between both binned rolls oriented to the North; the candidate auto-subtracts and disappears from the image. **Center-right:** Sum of the binned rolls in their original orientation; the candidate appears at two different position angles. **Right:** difference of the two binned rolls in their original orientations; the candidate appears in positive and negative at two different locations.

A standard format is currently being developed for direct imaging science product to enable consistent data interchange between different instruments (Choquet et al., these proceedings (9143-199)), in a specific FITS format including the final reduced image, contrast and SNR maps, and specific FITS keywords. The ALICE pipeline is currently being adapted to export high level science products in this format, that will be then exported to the MAST archive.

## 4. FIRST RESULTS

Although the pipeline has been constantly in development since the beginning of the ALICE project in 2012, we were still able to use it to reduce and partially analyze most of the NICMOS coronagraphic archive. All the targets observed with the F110W and F160W filters (about 70 % of the database) have been aligned in reference cubes, and at least preliminarily reduced to find interesting targets and targets with obvious circumstellar material. As soon as the pipeline is finalized with all features described herein, the high-level science products will be generated and delivered to the MAST archive.

### 4.1 Resolved Structures

The most unanticipated part of the ALICE project is its ability to reveal extended structures around a great number of targets. This can be attributed to two main reasons:

- The KLIP algorithm, which has good robustness to extended structures and is fast enough to explore very large parameter spaces in a few seconds, and find optimal parameters to reveal circumstellar material.
- The use of very large PSF libraries, which enables us to be very selective on the similarity between the references and the science image. When working with a 300-image library, discarding the 60 % less correlated with the science image still provides us with a cube of 120 reference PSF. In addition, since all of these are very good matches to the science image PSF, the algorithm is very efficient at removing PSF residuals at very small angular separations from the star, down to the photon noise level.

With the ALICE pipeline, we were able to easily re-image known disks and improve their detection especially at small separations. Fig. 9 present a few of these new reductions for four known disks. These new reductions enabled further analysis and characterization for some of these targets (Milli et al., submitted).

In addition to these previously known disks that are easily detected because of their brightness and/or angular size, we also revealed five disks in scattered-light for the first time, which were previously unseen in the NICMOS archive.<sup>27</sup> Only one of them (HD 202917) had already been marginally detected at low SNR in the visible with the HST Advanced Camera for Surveys (ACS). These discoveries increase by 21 % the number of debris disks imaged in scattered light. Three of these disks appear edge-on (HD 30447, HD 35841 and HD 141943) and the

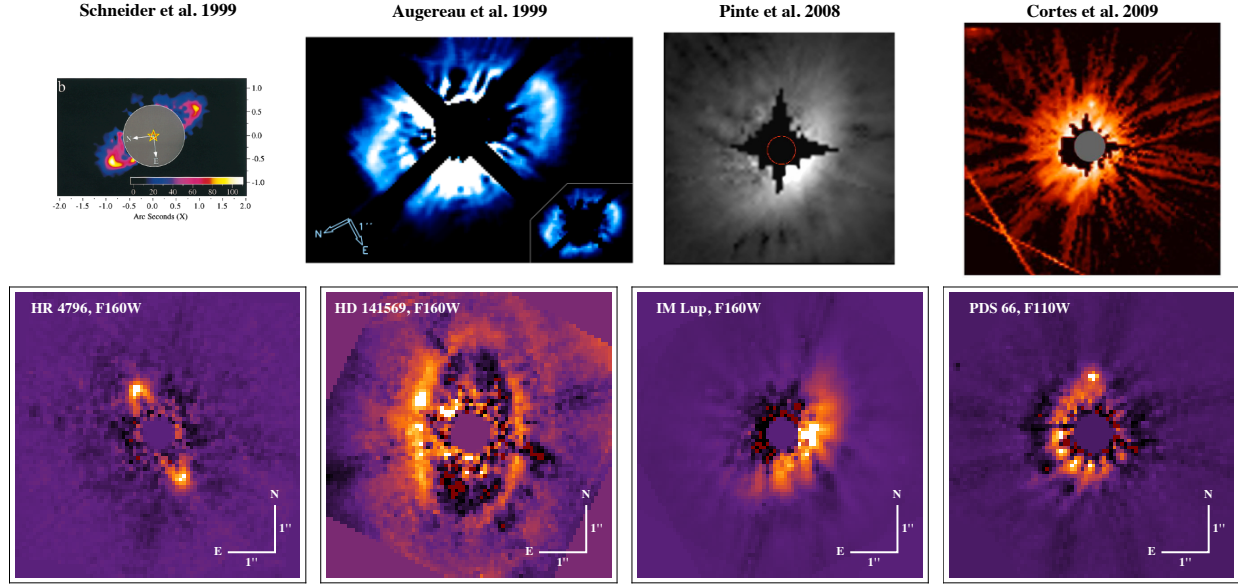


Figure 9. Known debris and protoplanetary disks imaged in scattered light with the coronagraphic mode of NICMOS. Top row: discovery images. Bottom row: reprocessing with the ALICE pipeline, with the exact same data-set. From left to right: HR 4796A (program 7233, F160W, discovery image from [23]), HD 141569 (program 7857, F160W, discovery image from [24]), IM Lup (program 10177, F160W, discovery image from [25]), PDS 66 (program 10527, F110W, discovery image from [26]).

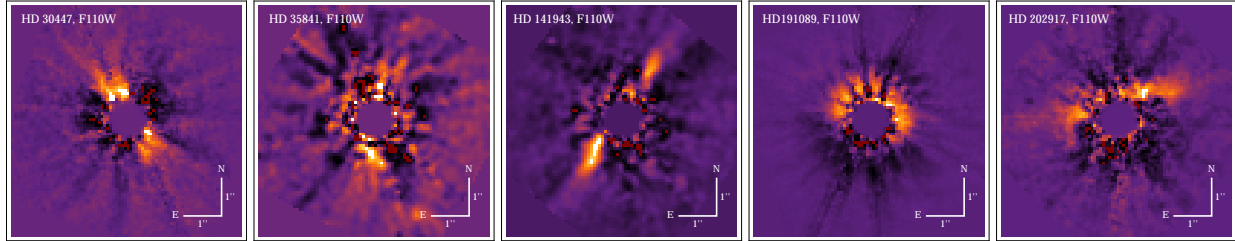


Figure 10. Five debris disks newly revealed in scattered light from NICMOS coronagraphic archive (filter F110W) with the ALICE pipeline.<sup>27</sup> From left to right: HD 30447 (F3V, 80 pc, 10-40 Myr in Columba association), HD 35841 (F3V, 96 pc, 10-40 Myr in Columba association), HD 141943 (G2V, 67 pc, 17-32 Myr), HD 191089 (F5V, 52 pc, 8-20 Myr in  $\beta$  Pictoris association), HD 202917 (G7V, 43 pc, 10-40 Myr in Tuc-Hor association).

other two are inclined (HD 191089 and HD 202917) (see Fig. 10). All five host stars are young (8 to 40 Myr), nearby (40 to 100 pc) main sequence stars (F and G type). In particular, HD 141943 is a close analog to the Sun at the age of terrestrial planet formation.

## 4.2 Point Sources

Additionally to these disks, we also imaged and characterized a large number of point sources. A number of them are very close binary systems with high magnitude difference, and were not previously detectable with classical techniques. We also found a dozen of substellar companion candidates, which have yet to be confirmed as co-moving with the host star using followup observations. Fig. 11 presents six candidates which are currently followed-up with HST-WFC3 to confirm their gravitational bound with the host star (PI: L. Pueyo). From stellar populations models, we estimated a probability of 98 % to have at least one gravitationally-bound companion among this sample. Followup observations of the other point source candidates that are bright enough to be observed with ground-based AO systems are also currently being conducted to confirm companionship with the host stars.

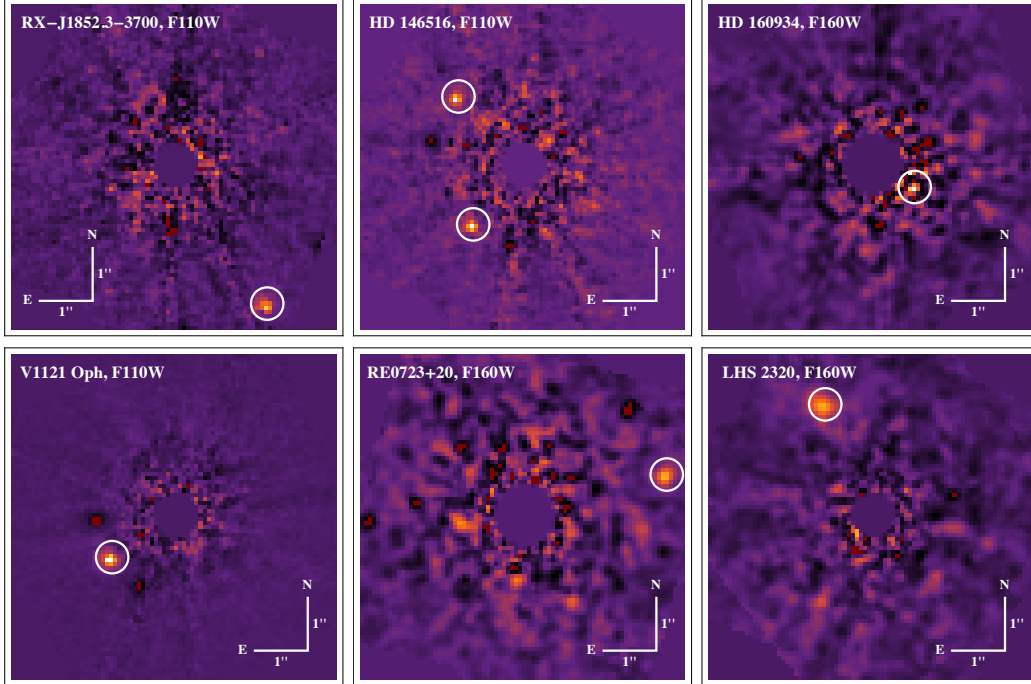


Figure 11. Six companion candidates discovered in the NICMOS coronagraphic archive with the ALICE pipeline and currently followed-up with the HST-WFC3 instrument (PI: L. Pueyo). From top-left to bottom-right: RX-J1852.3-3700 (K3, 130 pc, 3 Myr in Cha association), HD 146516 (G7V, 120 pc, 10 Myr, in Up. Sco. association), HD 160934 (K7Ve, 33 pc, 100 Myr in AB Dor. association), V1121 Oph (K4Ve, 125 pc, 1 Myr in  $\rho$  Oph association), RE0723+20 (K5Ve, 25 pc, 100 Myr in AB Dor. association), LHS 2320 (M5.0V, 22 pc, 100 Myr).

## 5. CONCLUSION

In this paper, we presented the ALICE project, which consists in a comprehensive and consistent reprocessing of the HST-NICMOS coronagraphic archive using advanced PSF subtraction algorithms. We briefly described the content of the archive, then provided a detailed description of the ALICE pipeline. Finally, we presented a summary of the first scientific results of this project obtained with a preliminary version of the pipeline, showing new images of debris disks in scattered light as well as unknown potential sub-stellar companions.

The pipeline is currently in its final development phase, consisting mostly in the optimization of the image alignment process and in the definition and implementation of a standard format for the high-level science products that will be generated for all the images in the archive. The entire archive will then be consistently reprocessed with the final version of the pipeline and the final products will be delivered to the MAST archive.

All of our discoveries are now included in followup observation campaigns. Follow-up observations of the point source candidates are being conducted with the HST-WFC3 for the six faintest targets, and with ground-based instruments (VLT-SINFONI, Keck, Palomar P1640) for the brighter ones others, to confirm their companionship and characterize their photometry and colors. In addition, an HST-STIS coronagraphic program is currently on-going to observe the five debris disks newly imaged with the ALICE pipeline, and will provide visible images of the disks by Fall 2014, with a resolution twice better than with NICMOS (PI: M. D. Perrin). These five disks are currently being analyzed by our team using the radiative transfer modeling MCFOST<sup>28</sup> to characterize their optical properties and composition.

## ACKNOWLEDGMENTS

This project was made possible by the Mikulski Archive for Space Telescopes (MAST) at STScI. Support was provided by NASA through grants HST-AR-12652.01 (PI: R. Soummer), HST-GO-11136.09-A (PI: D. Golimowski),

and by STScI Director’s Discretionary Research funds, from STScI, which is operated by AURA under NASA contract NAS5- 26555. The input images to ALICE processing are from the recalibrated NICMOS data products produced by the Legacy Archive project, “A Legacy Archive PSF Library And Circumstellar Environments (LAPLACE) Investigation,” (HST-AR- 11279, PI: G. Schneider). Pueyo was supported in part under contract with the California Institute of Technology (Caltech) funded by NASA through the Sagan Fellowship Program executed by the NASA Exoplanet Science Institute. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

## REFERENCES

- [1] Mayor, M. and Queloz, D., “A Jupiter-mass companion to a solar-type star,” *Nature* **378**, 355–359 (1995).
- [2] Collier Cameron, A., Bouchy, F., Hébrard, G., Maxted, P., Pollacco, D., Pont, F., Skillen, I., Smalley, B., Street, R. A., West, R. G., Wilson, D. M., Aigrain, S., Christian, D. J., Clarkson, W. I., Enoch, B., Evans, A., Fitzsimmons, A., Fleenor, M., Gillon, M., Haswell, C. A., Hebb, L., Hellier, C., Hodgkin, S. T., Horne, K., Irwin, J., Kane, S. R., Keenan, F. P., Loeillet, B., Lister, T. A., Mayor, M., Moutou, C., Norton, A. J., Osborne, J., Parley, N., Queloz, D., Ryans, R., Triaud, A. H. M. J., Udry, S., and Wheatley, P. J., “WASP-1b and WASP-2b: two new transiting exoplanets detected with SuperWASP and SOPHIE,” *MNRAS* **375**, 951–957 (2007).
- [3] Lovis, C., Ségransan, D., Mayor, M., Udry, S., Benz, W., Bertaux, J.-L., Bouchy, F., Correia, A. C. M., Laskar, J., Lo Curto, G., Mordasini, C., Pepe, F., Queloz, D., and Santos, N. C., “The HARPS search for southern extra-solar planets. XXVIII. Up to seven planets orbiting HD 10180: probing the architecture of low-mass planetary systems,” *A&A* **528**, A112 (2011).
- [4] Batalha, N. M., Rowe, J. F., Bryson, S. T., Barclay, T., Burke, C. J., Caldwell, D. A., Christiansen, J. L., Mullally, F., Thompson, S. E., Brown, T. M., Dupree, A. K., Fabrycky, D. C., Ford, E. B., Fortney, J. J., Gilliland, R. L., Isaacson, H., Latham, D. W., Marcy, G. W., Quinn, S. N., Ragozzine, D., Shporer, A., Borucki, W. J., Ciardi, D. R., Gautier, III, T. N., Haas, M. R., Jenkins, J. M., Koch, D. G., Lissauer, J. J., Rapin, W., Basri, G. S., Boss, A. P., Buchhave, L. A., Carter, J. A., Charbonneau, D., Christensen-Dalsgaard, J., Clarke, B. D., Cochran, W. D., Demory, B.-O., Desert, J.-M., Devore, E., Doyle, L. R., Esquerdo, G. A., Everett, M., Fressin, F., Geary, J. C., Girouard, F. R., Gould, A., Hall, J. R., Holman, M. J., Howard, A. W., Howell, S. B., Ibrahim, K. A., Kinemuchi, K., Kjeldsen, H., Klaus, T. C., Li, J., Lucas, P. W., Meibom, S., Morris, R. L., Prša, A., Quintana, E., Sanderfer, D. T., Sasselov, D., Seader, S. E., Smith, J. C., Steffen, J. H., Still, M., Stumpe, M. C., Tarter, J. C., Tenenbaum, P., Torres, G., Twicken, J. D., Uddin, K., Van Cleve, J., Walkowicz, L., and Welsh, W. F., “Planetary Candidates Observed by Kepler. III. Analysis of the First 16 Months of Data,” *ApJS* **204**, 24 (2013).
- [5] Soummer, R., Brendan Hagan, J., Pueyo, L., Thormann, A., Rajan, A., and Marois, C., “Orbital Motion of HR 8799 b, c, d Using Hubble Space Telescope Data from 1998: Constraints on Inclination, Eccentricity, and Stability,” *ApJ* **741**, 55 (2011).
- [6] Konopacky, Q. M., Barman, T. S., Macintosh, B. A., and Marois, C., “Detection of Carbon Monoxide and Water Absorption Lines in an Exoplanet Atmosphere,” *Science* **339**, 1398–1401 (2013).
- [7] Oppenheimer, B. R., Baranec, C., Beichman, C., Brenner, D., Burruss, R., Cady, E., Crepp, J. R., Dekany, R., Fergus, R., Hale, D., Hillenbrand, L., Hinkley, S., Hogg, D. W., King, D., Ligon, E. R., Lockhart, T., Nilsson, R., Parry, I. R., Pueyo, L., Rice, E., Roberts, J. E., Roberts, Jr., L. C., Shao, M., Sivaramakrishnan, A., Soummer, R., Truong, T., Vasisht, G., Veicht, A., Vescelus, F., Wallace, J. K., Zhai, C., and Zimmerman, N., “Reconnaissance of the HR 8799 Exosolar System. I. Near-infrared Spectroscopy,” *ApJ* **768**, 24 (2013).
- [8] Hinkley, S., Pueyo, L., Faherty, J. K., Oppenheimer, B. R., Mamajek, E. E., Kraus, A. L., Rice, E. L., Ireland, M. J., David, T., Hillenbrand, L. A., Vasisht, G., Cady, E., Brenner, D., Veicht, A., Nilsson, R., Zimmerman, N., Parry, I. R., Beichman, C., Dekany, R., Roberts, J. E., Roberts, Jr., L. C., Baranec, C., Crepp, J. R., Burruss, R., Wallace, J. K., King, D., Zhai, C., Lockhart, T., Shao, M., Soummer, R., Sivaramakrishnan, A., and Wilson, L. A., “The  $\kappa$  Andromedae System: New Constraints on the Companion Mass, System Age, and Further Multiplicity,” *ApJ* **779**, 153 (2013).
- [9] Golimowski, D. A., Krist, J. E., Stapelfeldt, K. R., Chen, C. H., Ardila, D. R., Bryden, G., Clampin, M., Ford, H. C., Illingworth, G. D., Plavchan, P., Rieke, G. H., and Su, K. Y. L., “Hubble and Spitzer Space Telescope Observations of the Debris Disk around the nearby K Dwarf HD 92945,” *AJ* **142**, 30 (July 2011).

- [10] Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., Lafrenière, D., and Doyon, R., “Direct Imaging of Multiple Planets Orbiting the Star HR 8799,” *Science* **322**, 1348– (2008).
- [11] Kalas, P., Graham, J. R., Chiang, E., Fitzgerald, M. P., Clampin, M., Kite, E. S., Stapelfeldt, K., Marois, C., and Krist, J., “Optical Images of an Exosolar Planet 25 Light-Years from Earth,” *Science* **322**, 1345– (2008).
- [12] Lagrange, A.-M., Bonnefoy, M., Chauvin, G., Apai, D., Ehrenreich, D., Boccaletti, A., Gratadour, D., Rouan, D., Mouillet, D., Lacour, S., and Kasper, M., “A Giant Planet Imaged in the Disk of the Young Star  $\beta$  Pictoris,” *Science* **329**, 57– (2010).
- [13] Krist, J. E., Golimowski, D. A., Schroeder, D. J., and Henry, T. J., “Characterization and Subtraction of Well-Exposed HST/NICMOS Camera 2 Point-Spread Functions for a Survey of Very Low Mass Companions to Nearby Stars,” *PASP* **110**, 1046–1058 (1998).
- [14] Lallo, M. D., Makidon, R. B., Casertano, S., and Krist, J. E., “Temporal optical behavior of HST: focus, coma, and astigmatism history,” in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **6270** (July 2006).
- [15] Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., and Artigau, É., “A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging,” *ApJ* **660**, 770–780 (2007).
- [16] Soummer, R., Pueyo, L., and Larkin, J., “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages,” *ApJ* **755**, L28 (2012).
- [17] Amara, A. and Quanz, S. P., “PYNPOINT: an image processing package for finding exoplanets,” *MNRAS* **427**, 948–955 (2012).
- [18] Lafrenière, D., Marois, C., Doyon, R., and Barman, T., “HST/NICMOS Detection of HR 8799 b in 1998,” *ApJ* **694**, L148–L152 (2009).
- [19] Schneider, G., Silverstone, M. D., Stobie, E., Rhee, J. H., and Hines, D. C., “NICMOS Coronagraphy: Recalibration and the NICMOS Legacy Archive PSF Library,” in [*Hubble after SM4. Preparing JWST*], (2010).
- [20] Krist, J. E., Hook, R. N., and Stoehr, F., “20 years of Hubble Space Telescope optical modeling using Tiny Tim,” in [*Proc. SPIE*], *Proc. SPIE* **8127** (2011).
- [21] Gratadour, D., Mugnier, L. M., and Rouan, D., “Sub-pixel image registration with a maximum likelihood estimator. Application to the first adaptive optics observations of Arp 220 in the L’ band,” *A&A* **443**, 357–365 (2005).
- [22] Lowrance, P. J., Becklin, E. E., Schneider, G., Kirkpatrick, J. D., Weinberger, A. J., Zuckerman, B., Dumas, C., Beuzit, J.-L., Plait, P., Malumuth, E., Heap, S., Terrile, R. J., and Hines, D. C., “An Infrared Coronagraphic Survey for Substellar Companions,” *AJ* **130**, 1845–1861 (2005).
- [23] Schneider, G., Smith, B. A., Becklin, E. E., Koerner, D. W., Meier, R., Hines, D. C., Lowrance, P. J., Terrile, R. J., Thompson, R. I., and Rieke, M., “NICMOS Imaging of the HR 4796A Circumstellar Disk,” *ApJ* **513**, L127–L130 (1999).
- [24] Augereau, J. C., Lagrange, A. M., Mouillet, D., and Ménard, F., “HST/NICMOS2 observations of the HD 141569 A circumstellar disk,” *A&A* **350**, L51–L54 (1999).
- [25] Pinte, C., Padgett, D. L., Ménard, F., Stapelfeldt, K. R., Schneider, G., Olofsson, J., Panić, O., Augereau, J. C., Duchêne, G., Krist, J., Pontoppidan, K., Perrin, M. D., Grady, C. A., Kessler-Silacci, J., van Dishoeck, E. F., Lommen, D., Silverstone, M., Hines, D. C., Wolf, S., Blake, G. A., Henning, T., and Stecklum, B., “Probing dust grain evolution in IM Lupi’s circumstellar disc. Multi-wavelength observations and modelling of the dust disc,” *A&A* **489**, 633–650 (2008).
- [26] Cortes, S. R., Meyer, M. R., Carpenter, J. M., Pascucci, I., Schneider, G., Wong, T., and Hines, D. C., “Grain Growth and Global Structure of the Protoplanetary Disk Associated with the Mature Classical T Tauri Star, PDS 66,” *ApJ* **697**, 1305–1315 (2009).
- [27] Soummer, R., Perrin, M. D., Pueyo, L., Choquet, É., Chen, C., Golimowski, D. A., Brendan Hagan, J., Mittal, T., Moerchen, M., N’Diaye, M., Rajan, A., Wolff, S., Debes, J., Hines, D. C., and Schneider, G., “Five Debris Disks Newly Revealed in Scattered Light from the Hubble Space Telescope NICMOS Archive,” *ApJ* **786**, L23 (2014).



- [28] Pinte, C., Ménard, F., Duchêne, G., and Bastien, P., “Monte Carlo radiative transfer in protoplanetary disks,” *A&A* **459**, 797–804 (Dec. 2006).